

Conjuntos y Sistemas Difusos

(Lógica Difusa y Aplicaciones)

9. Bases de Datos Relacionales Difusas (BDRD):

Modelos Teóricos, Álgebra y Cálculo Relacional Difuso



E.T.S.I. Informática

J. Galindo Gómez

Modelo Relacional de BD

- **Fue propuesto en 1970 por el Dr. E.F. Codd de IBM**, como un intento de simplificar la **estructura de las bases de datos**, demasiado compleja en otros modelos implementados en distintos **SGBD** (Sistema Gestor de Bases de Datos, DBMS, *DataBase Management System*).
 - El sistema se generalizó rápidamente, aunque no todos los sistemas que se llamaban “relacionales” lo eran. Por eso, Codd publicó en 1986 una lista de 12 reglas que deberían cumplir los SGBD relacionales.
- **Elementos del modelo relacional:** { Una visión moderna y completa está en (Elmasri, Navathe, 2000).
 - **Estructuras de datos:** Tablas o relaciones (atributos, dominios, tuplas...).
 - **Integridad de los Datos:** Claves o llaves (candidata, primaria y externa).
 - **Regla de Identidad:** No se admite NULL en la clave primaria.
 - **Regla de Integridad Referencial:** La clave externa debe existir (como primaria).
 - **Lenguaje de Definición de Datos, DDL** (*Data Definition Language*): Para crear, borrar y modificar objetos de la BD (tablas, vistas, índices).
 - **Lenguaje de Manipulación de Datos, DML** (*Data Manipulation Language*): Para consultar, insertar y modificar datos de la BD.
 - Dos niveles de lenguajes formales de consulta:
Álgebra Relacional y **Cálculo Relacional** (Codd, 1972).

Representando Imprecisión en BD

- Existen **varios modelos** para representar imprecisión en BD:
 - Cada uno tiene sus **ventajas, sus desventajas y sus limitaciones**.
 - El objetivo global va más allá de simplemente representar la información imprecisa, sino que se requieren **modificar también las operaciones que se efectúan sobre las relaciones** (consultas difusas o flexibles, condiciones, inserción y modificación de datos difusos...).
 - También se incluye la **consulta flexible o difusa en bases de datos clásicas** (sin más imprecisión que los NULL del modelo relacional según el SGBD empleado).
 - Aquí veremos algunos de los modelos más importantes:
 - Modelos SIN Lógica Difusa:** Ideas de Codd, Date, Grant, Wong...
 - Modelos de BDRD:** Modelo de Prade-Testemale, Modelo de Umano-Fukami, Modelo de Buckles-Petry, Modelo de Zemankova-Kandel y el Modelo GEFRED de Medina-Pons-Vila.
 - Bibliografía general:** (Petry, 1996; Medina, 1994; Galindo, 1999).
 - Las dos últimas referencias pueden conseguirse por internet.
- Es un tema **muy estudiado**, pero no por ello está cerrado, y constantemente surgen ideas para solucionar problemas particulares o añadir ventajas adicionales con algún objetivo.

3

Imprecisión en BD, SIN Lógica Difusa

- Aproximación de Codd:** (Codd, 1979, 1986, 1987, 1990).
 - Codd introdujo el valor **NULL**, para indicar que el valor de ese atributo era **desconocido** y, por tanto, era posible cualquier valor del dominio.
 - Comparaciones:** Se usa una **lógica tri-valuada** de forma que cualquier comparación con el valor **NULL** no genera ni Verdad (T) ni Falso (F), sino **quizás (m, maybe)**:

NOT		AND	T	m	F	OR	T	m	F
T	F	T	T	m	F	T	T	T	T
m	m	m	m	m	F	m	T	m	m
F	T	F	F	F	F	F	T	m	F

Esta es la lógica usada por el **SGBD Oracle**, donde m es interpretado como "Unknown".

- Posteriormente distinguió entre **dos marcas (A, I)**, generando una **lógica tetra-valuada**, (los A e I surgen de comparaciones con esas marcas):
 - A-marca:** Valor ausente o desconocido, pero aplicable.
 - I-marca:** Valor no aplicable (Ej.: Matrícula del coche de alguien sin coche).

NOT		AND	T	A	I	F	OR	T	A	I	F
T	F	T	T	A	I	F	T	T	T	T	T
A	A	A	A	A	I	F	A	T	A	A	A
I	I	I	I	I	I	F	I	T	A	I	F
F	T	F	F	F	F	F	F	T	A	F	F

4

Imprecisión en BD, SI N Lógica Difusa

- **Esquema de Valores por Defecto, de Date** (1982):
 - Date fue reacio a la utilización de los valores **NULL** porque, a su juicio no estaba bien definido su comportamiento.
 - Propuso que al definir cada dominio de cada atributo se puede asignar un **valor por defecto** (similar a la cláusula **DEFAULT** de **SQL**).
- **Rangos en Valores, de Grant** (1980):
 - Grant extendió el modelo relacional para permitir almacenar dos valores en un atributo **[a,b]** con el significado de un **rango de valores posibles**.
 - También acepta el **valor NULL** si no se tiene ninguna información.
 - En su modelo **se permiten tuplas repetidas**, ya que aunque tengan el mismo rango pueden tener valores distintos.
 - **Los operadores relacionales son redefinidos en dos versiones:**
 - **True (auténtica):** Ejemplo: $[a,b] <_T [n,m]$ es verdad si $b < n$.
 - **Maybe (posible):** Ejemplo: $[a,b] <_M [n,m]$ es verdad si $a < m$.
 - **El resultado de una consulta se divide en tres partes:** Tuplas que **seguro que pertenecen al resultado**, tuplas que **posiblemente pertenecen al resultado** y tuplas que **seguro que no pertenecen a él**.

5

Imprecisión en BD, SI N Lógica Difusa

- **Bases de Datos Estadísticas y Probabilísticas:**
 - **Barbara et al. (1992)** definieron un modelo de BD en el que cada atributo puede tener asociada una distribución de probabilidad
 - La suma de todas las probabilidades debe ser igual a 1.
 - Resulta difícil calcular la probabilidad para todos los valores del dominio, por lo que meten las probabilidades conocidas y las desconocidas las meten en las llamadas probabilidades perdidas cuya probabilidad será 1 menos la suma de las probabilidades conocidas.
 - Cálculo de probabilidades:
 - Introducidas por el usuario.
 - Calculadas a partir de un grupo de ejemplos (encuestas...).
 - En las consultas podría establecerse un umbral de probabilidad mínimo.
 - **Otros modelos:**
 - Unos asocian la probabilidad a cada tupla según la probabilidad de que ésta pertenezca a la relación (Cavallo, Pittarelli, 1987).
 - Wong primero y luego junto con Shosnani (1985) publicaron un sistema de consulta en la que ésta es vista como un experimento estadístico cuyo objetivo es minimizar el error estadístico.
 - Algunos trabajos se centran en consultas imprecisas con probabilidad (Fuhr, 1990).

6

Modelos de BDRD: Básicos

- **Modelos Básicos:**

- Consisten en añadir un “grado” en el intervalo $[0,1]$.
- El significado de este grado puede variar, pero es fundamental y determinante en los procesos de consulta sobre este tipo de tablas.
- Formas de añadir ese grado difuso:
 - **Grado en cada tupla de la relación:** El grado difuso pertenece a toda la tupla de la relación, por lo que lo que se “difumina” es la relación propiamente dicha (Mouabdib, 1994; Tahani, 1977; Umano, Fukami, 1994).
 - No permite información imprecisa sobre un atributo en particular.
 - **Grado en cada valor de cierto atributo:** Mide el grado de “difuminado” de ese valor concreto del atributo en la tupla a la que pertenezca (Bosc et al., 1987; Medina et al., 1994).
 - Tampoco permiten información imprecisa distinta de un único grado difuso. Por ejemplo, almacenar que cierta persona es “Joven”.
 - **Grado en un conjunto de valores:** El grado afecta a varios atributos midiendo, por ejemplo, la incertidumbre que hay en ellos.

7

Modelos de BDRD: Básicos

- **Algunos Significados de los “Grados” difusos:**

- **Grado de Pertenencia, de Incertidumbre o de Posibilidad:** Son tres significados distintos pero similares.
 - El grado de Pertenencia mide la pertenencia a un conjunto difuso.
 - El grado de Incertidumbre mide en qué medida es cierto el dato.
 - El grado de Posibilidad mide lo posible que es el dato.
 - Refs.: (Mouabdib, 1994; Umano, Fukami, 1994).
- **Nivel de Dependencia** entre los atributos de la relación.
 - Refs.: (Baldwin, 1983).
- **Grado de Cumplimiento de una condición o de una propiedad.**
 - Puede verse también como grado de Pertenencia al conjunto difuso de los elementos que cumplen tal condición o propiedad.
 - Refs.: (Tahani, 1977; Medina et al., 1994; Bosc et al., 1997; Galindo, 1999).
 - Puede almacenarse o simplemente calcularse a partir de una consulta difusa (Galindo et al., 1998a).
- **Grado de Importancia:** En una relación, distintos objetos pueden tener diferente nivel de importancia y esa diferencia puede ser útil considerar para algunos propósitos.
 - Refs.: (Mouabdib, 1994; Bosc et al., 1997).

8

Modelos de BDRD: Atributos Difusos

- **Valores Difusos en los Atributos de las Relaciones:**

- ¿Por qué incorporar imprecisión en los atributos de las BD?
 - Las bases de datos tradicionales son muy limitadas: No permiten ni almacenar ni tratar con datos imprecisos.
 - Las personas manejamos datos imprecisos muy a menudo y muy eficientemente.

- Ejemplo:

JUGADOR	EQUIPO	ALTURA	CALIDAD
J1	Córdoba	Bajo	Muy Bueno
J2	Granada	Alto	Bueno
J3	Sevilla	#210	Regular
J4	Almería	Normal	Malo

- Sería ideal tener mecanismos en los SGBD ya existentes, para:
 - Almacenar imprecisión en bases de datos.
 - Poder consultar la base de datos a través de consultas flexibles.
- Es preferible utilizar SGBD ya existentes para poder reutilizar el resto del SGBD y para poder implantar BDRD en aquellos sistemas en los que ese SGBD esté implantado.

9

Modelos de BDRD: Objetivos

- **Objetivos Principales de las BDRD:**

- 1. **Almacenar Imprecisión**, la información que tengamos de un atributo particular de un objeto, aunque esta información no sea el valor exacto.
 - Suelen usar **Etiquetas Lingüísticas** con alguna definición asociada (por ejemplo, un conjunto difuso visto como una “**Distribución de Posibilidad**”), o sin ninguna definición asociada (“**escalares**” con una relación de similitud definida entre ellos).
- 2. **Operar con esa información** de forma coherente (especialmente en las operaciones de **consulta**).
 - Muchos autores estudian la consulta difusa en BD clásicas: (Tahani, 1977; Bosc et al., 1988, 1994, 1995; Wong, 1990; Galindo et. al., 1998a).

- **Modelos más importantes publicados:**

- **Modelo de Buckles-Petry:** (Buckles, Petry, 1982, 1984).
- **Modelo de Umano-Fukami:** (Umano, 1982; Umano, Fukami, 1994).
- **Modelo de Prade-Testemale:** (Prade, Testemale, 1984, 1987).
- **Modelo de Zemankova-Kandel:** (Zemankova, Kandel, 1984, 1985).
- **Modelo GEFRED de Medina-Pons-Vila:** (Medina et al., 1994).

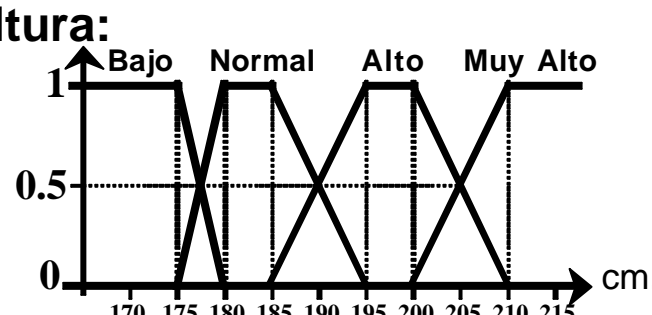
10

Modelos de BDRD: Buckles-Petry

- **Modelo de Buckles-Petry:** Fue el primer modelo de BDRD que utiliza relaciones de similitud.
 - Un atributo puede tomar “**varios**” valores de su dominio, cuyo significado es disyuntivo (el valor real es uno y sólo uno de los valores especificados pero ignoramos cual de ellos es).
 - En cada dominio se define una “**Relación de Similitud**” que mide el parecido entre cada dos elementos del dominio D .
 - Suele **normalizarse** en el intervalo $[0,1]$, correspondiendo el 0 a “totalmente diferentes” y el 1 a “extremadamente parecidos o iguales”.
 - **Relación de Similitud:** Es una función $S: D \times D \rightarrow [0,1]$
 - Cumple las propiedades Reflexiva, Simétrica y Transitiva.
 - La propiedad transitiva hace que se puedan construir “**Clases de Equivalencia**” para un determinado umbral γ , de forma que los elementos de una misma clase sean indistinguibles para el grado γ .
 - Otros autores (Shenoi, Melton, 1990) usan “**Relaciones de Proximidad**”, las cuales no tienen que cumplir la propiedad transitiva.
 - También puede resultar útil evitar la propiedad Simétrica.

11

Modelos de BDRD: Umano-Fukami

- **Modelo de Umano-Fukami:** Es un modelo de los llamados “posibilísticos” porque usan “distribuciones de posibilidad” para modelar la información conocida sobre el valor de un atributo.
 - Sobre el dominio X de un atributo, la distribución de posibilidad A indica que la posibilidad de que el atributo tome el valor $x \in X$ es $A(x)$.
 - **Ejemplos:** Cuatro etiquetas lingüísticas asociadas a otras cuatro distribuciones de posibilidad para el atributo “Altura”.
 
 - Un valor x posible no indica que sea cierto (aunque su posibilidad sea 1).
 - Pueden **compararse** distribuciones (medidas de necesidad/posibilidad...).
 - **Valores especiales de distribuciones de posibilidad:**
 - **Unknown:** Desconocido, pero aplicable: $A(x) = 1, \forall x \in X$.
 - **Undefined:** Valor no aplicable o sin sentido: $A(x) = 0, \forall x \in X$.
 - **Null:** Ignorancia total, no se sabe si es o no aplicable.
- Distribución de posibilidad: $\text{Null} = \{1/\text{Unknown}, 1/\text{Undefined}\}$.

12

Modelos de BDRD: Umano-Fukami

- El modelo de Umano-Fukami permite además asociar a cada tupla una **Distribución de Posibilidad** en el intervalo $[0,1]$.
 - Esta distribución mide el **Grado de Pertenencia** difusa de cada tupla a la relación: La tabla es como un conjunto difuso de tuplas.
 - La distribución de posibilidad en el intervalo $[0,1]$, consigue **mayor expresividad**. Por ejemplo, se puede expresar que cierta tupla pertenece a la relación con grado “aproximadamente 0.6”.
 - Por supuesto, permite también almacenar un **Grado Simple** en el intervalo $[0,1]$. Por ejemplo, el valor 0.8 corresponde a la distribución de posibilidad $\{1/0.8\}$.
- Umano y Fukami definen también las operaciones básicas del **Álgebra Relacional** (Unión, Diferencia, Producto Cartesiano, Proyección y Selección), así como otras operaciones (Intersección, Reunión y División).
- Si hay que operar con distribuciones de posibilidad utiliza el **Principio de Extensión**.
- Otros autores sugieren que en las **consultas** se tome para cada tupla el **mínimo** entre el grado de la relación original y el grado de cumplimiento de la condición de la consulta.

13

Modelos de BDRD: Prade-Testemale

- **Modelo de Prade-Testemale:** Modelo posibilístico que añade un elemento “e” al dominio de todos los atributos, que representa el caso en el que el atributo no sea aplicable. **Comparación de representación:**

Información	Modelo Prade-Testemale	Modelo Umano-Fukami
Sabemos el dato y este es <i>crisp</i> : c	$A(e)=0; A(c)=1;$ $A(x)=0, "x \hat{\in} X, x \neq c;$	$A(x) = \{1/c\};$
Desconocida (pero aplicable)	$A(e)=0;$ $A(x)=1, "x \hat{\in} X;$	Unknown
No aplicable	$A(e)=1; A(x)=0, "x \hat{\in} X;$	Undefined
Ignorancia total	$A(x)=1, "x \hat{\in} X \setminus \{e\};$	Null
Rango [m,n]	$A(x)=0, "x \hat{\in} [m,n] \text{ o } x = e;$ $A(x)=1, "x \hat{\in} [m,n];$	$A(x)=0, "x \hat{\in} [m,n];$ $A(x)=1, "x \hat{\in} [m,n];$
Distribución de Posibilidad B	$A(e)=0;$ $A(x)=B(x), "x \hat{\in} X;$	$A(x)=B(x), "x \hat{\in} X;$
Posibilidad de que no sea aplicable es 1 y si lo es vale B.	$A(e)=1;$ $A(x)=B(x), "x \hat{\in} X;$	No representable

14

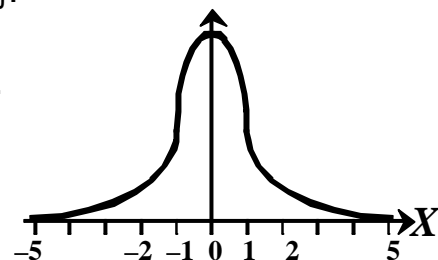
Modelos de BDRD: Zemankova-Kandel

- **Modelo de Zemankova-Kandel:** Es otro modelo posibilístico.
 - No usa las medidas de Posibilidad y Necesidad para hacer las comparaciones entre dos distribuciones de posibilidad, sino que usa las siguientes dos medidas para comparar el atributo A con el conjunto difusos F :
 - **Medida de Posibilidad:** $p_A(F) = \sup_{x \in X} \{F(x) \cdot A_i(x)\}$;
 - **Medida de Certeza:** $c_A(F) = \max\{0, \inf_{x \in X} \{F(x) \cdot A_i(x)\}\}$;
 - Las medidas de Posibilidad/Necesidad son más coherentes:
 - La medida de certeza no tiene una interpretación clara.
 - No están relacionadas entre sí, como sí lo están las medidas de Posibilidad/Necesidad: $N(X) = 1 - P(\neg X)$.
 - En una consulta aparecerán esos dos valores, para cada tupla (2 columnas extra).
 - Definen unos **comparadores difusos** poco intuitivos: “aproximadamente igual”, “mayor que” y “menor que”.

15

Modelos de BDRD: Modelo GEFRED

- **Modelo de GEFRED de Medina-Pons-Vila:** Es una síntesis ecléctica de los otros modelos.
 - **Dominio Difuso Generalizado:** Es un modelo posibilístico, por lo que en los dominios admite distribuciones de posibilidad, pero también incluye el caso en el que el dominio subyacente no sea numérico sino escalares de cualquier tipo.
 - **Ejemplos:**
 - Distribución de posibilidad sobre un atributo difuso con dominio subyacente no ordenado como “Color del Pelo”: {1/Castaño, 0.7/Pelirrojo, 0.4/Rubio}.
 - Números difusos de cualquier tipo: Trapecios, triángulos, gaussianos...



- Incluye los valores **Unknown**, **Undefined** y **Null** con el mismo sentido que en el modelo de **Umano-Fukami**.

16

Modelos de BDRD: Modelo GEFRED

• Modelo de GEFRED de Medina-Pons-Vila:

- **Relación Difusa Generalizada:** Es una relación cuyos atributos tienen un Dominio Difuso Generalizado.
 - Además, a cada atributo A_j es posible asociarle un “**atributo de compatibilidad**” C_j donde almacenar un grado:
 - Se obtiene como consecuencia de los procesos de manipulación de los datos de esa relación.
 - Expresa el grado con el que ese valor ha satisfecho la operación realizada sobre él.
 - Se compone de:
 - **Cabeza H:** Nombre de cada uno de los n atributos, sus dominios y sus atributos de compatibilidad (opcional)
 - **Cuerpo B:** Incluye los valores de m tuplas: $i = 1, 2, \dots, m$.

$$R = \begin{cases} H = \{(A_1 : D_1[, C_1]), \dots, (A_n : D_n[, C_n])\} \\ B = \{(A_1 : \tilde{d}_{i1}[, c_{i1}]), \dots, (A_n : \tilde{d}_{in}[, c_{in}])\} \end{cases}$$

17

Modelo GEFRED: Comparadores

- **Comparador Difuso Generalizado:** GEFRED define un tipo de comparador general basado en cualquier comparador clásico existente ($=$, $>$, $<$...), pero no concreta la definición de cada uno.
 - El único requisito que establece es que el Comparador Difuso debe **respetar los resultados de los comparadores clásicos** cuando se comparan distribuciones de posibilidad que expresan valores *crisp* (como $1/x$ con $x \in X$).
 - Igualdad Difusa o Aproximadamente Igual: La igualdad es el comparador difuso más importante y más empleado.
 - Suelen usarse las siguientes medidas:
 - de **Posibilidad** : $\text{Poss}(A, B) = \sup_{x \in X} \{\min(A(x), B(x))\}$
 - » Posibilidad de que A y B sean iguales
 - de **Necesidad** : $\text{Nec}(A, B) = \inf_{x \in X} \{\max(A(x), 1 - B(x))\}$
 - » Necesidad de que B sea igual a A (no a la inversa).
 - GEFRED no propone otros comparadores difusos.
 - Pueden encontrarse la definición de otros comparadores (Mayor difuso, Mayor o Igual difuso, Mucho mayor difuso) en su versión de posibilidad y necesidad, así como la comparación de distribuciones de posibilidad sobre dominios subyacentes no ordenado en (Galindo et al, 1998b; Galindo et al. 2000b).

18

Modelo GEFRED: Clave Primaria

- **Indistinguibilidad/Redundancia**: En una relación no pueden existir tuplas repetidas.
 - Si se diferencian sólo en atributos difusos: ¿Cómo discernir si son o no la misma tupla?
 - **Prade-Testemale** (1987) proponen que dos distribuciones de posibilidad A y B son aproximadamente iguales (indistinguibles) si
$$\sup_{x \in X} |A(x) - B(x)| \leq \varepsilon$$
donde ε es un umbral que depende del dominio X que se considere.
 - Dos tuplas serán consideradas iguales (redundantes) si son aproximadamente iguales en todos sus componentes.
 - **Problemas**:
 - Ese criterio no cumple la propiedad transitiva por lo que no se pueden crear clases de equivalencia.
 - Dados dos valores redundantes ¿qué valor almacenamos en la BD? Lo más intuitivo es almacenar la unión de ambos.
 - También pueden considerarse dos valores como iguales si uno incluye al otro, de forma que se elimina el valor que está incluido en el otro.
 - Pero, ¿qué hacemos si en una tupla se dan inclusiones opuestas?
- **Clave Primaria**: GEFRED es conservador, exigiendo atributos crisp o difusos con algún criterio de redundancia (como ser exactamente iguales).

19

Modelo GEFRED: Álgebra Relac. Difusa

- **GEFRED** se centra en cómo calcular los grados de compatibilidad de la relación difusa resultante y no en cómo calcular los valores de las tuplas de dicha relación.
- **Unión Difusa**: $R \dot{\cup} S$
 - Requiere que las dos relaciones difusas, R y S , sean **compatibles respecto de la unión** (con el mismo número y tipo de atributos).
 - **Resultado** de la unión de ellas: Es otra relación difusa con la misma cabecera y con las tuplas de ambas relaciones quitando redundancias.
 - Cuando halla redundancia se toma como **atributo de compatibilidad** el valor **máximo** de ambos atributos de compatibilidad (si existen ambos y si no se toma el que exista o ninguno, si no existe ninguno).
- **Intersección Difusa**: $R \dot{\cap} S$
 - **Resultado**: Es otra relación difusa con la misma cabecera y conteniendo las tuplas que existen a la vez en ambas relaciones.
 - Similar a la unión pero en este caso si existen los dos atributos de compatibilidad de las dos relaciones, el **atributo de compatibilidad** de la relación resultante será el valor **mínimo** de ambos atributos de compatibilidad.
 - **Propiedades de la Unión/Intersección**: Conmutativa y Asociativa.

20

Modelo GEFRED: Álgebra Relac. Difusa

- **Diferencia Difusa: $R - S$**
 - **Resultado:** Es otra relación difusa con la misma cabecera y con las tuplas que existen en la primera relación y no existen en la segunda.
 - El **Grado de Compatibilidad** para cada tupla del resultado se calcula como: $\min\{C_R, 1 - C_S\}$, donde C_R y C_S son los grados de compatibilidad de R y S respectivamente.
 - Si no existe C_R se supone que su valor es 1 y si no existe C_S se supone 0.
 - Esta operación resulta especialmente interesante para muchos tipos de consultas: GEFRED no define cómo calcular el resultado, sino sólo cómo calcular los **grados** del resultado.
- **Producto Cartesiano Difuso: $R \times S$**
 - **Resultado:** Es otra relación difusa teniendo como cabecera la unión de las cabeceras de R y S , y con cuerpo el producto cartesiano de los cuerpos de R y S .
- **Proyección Difusa: $P_X(R)$**
 - **Resultado:** Es otra relación difusa con cabecera los atributos de X , y con cuerpo los valores de esos atributos en las tuplas de R , eliminando posibles tuplas redundantes según el criterio que se adopte. 21

Modelo GEFRED: Álgebra Relac. Difusa

- **Selección Difusa: $s_C(R)$**
 - **Resultado:** Es otra relación difusa con la misma cabecera y con las tuplas que cumplen la condición C , con un umbral mínimo γ establecido en la misma condición.
 - El **grado de compatibilidad** para cada atributo del resultado es el grado con el que cada atributo satisface la condición C .
 - Aunque el modelo no lo indica, puede añadirse un grado adicional para toda la tupla que indique el grado con el que la condición ha sido satisfecha por toda la tupla y no por cada atributo individualmente.
 - Tampoco se aclara qué hacer cuando los atributos de la condición C tienen atributos de compatibilidad de operaciones anteriores.
- **Reunión Difusa: $R \bowtie_C S = s_C(R \times S)$**
 - **Resultado:** Es una selección con la condición C sobre el producto cartesiano $R \times S$.
- **División Difusa: $R \div S$**
 - Es una operación especialmente interesante, útil y compleja, que GEFRED no incluye en su definición original. Por eso, la estudiaremos aparte comparando diversos modelos de división relacional difusa. 22

CÁLCULO RELACIONAL DIFUSO

- **Objetivos:**

- Definir un lenguaje de consulta para BDRD, basado en el Cálculo Relacional de Codd (1972), que permita usar Cuantificadores Difusos.
- **Método para calcular los grados de cumplimiento** de la condición (función Δ): Esto es fundamental en BDRD, pues una de las ventajas básicas es considerar las tuplas resultantes como un conjunto difuso.

- **Existen distintas propuestas:**

- Buckles et al. (1989), propusieron un cálculo relacional difuso para su modelo de bases de datos relacionales.
 - No permiten establecer un umbral para cada átomo (condición simple), sino para cada atributo.
 - No permiten cuantificadores difusos (sólo \forall y \exists).
 - No definen un método claro para calcular los grados de cumplimiento, ya que su modelo no tiene los atributos de compatibilidad del modelo GEFRED.
- Nosotros veremos a continuación una definición adaptada para el modelo GEFRED (Galindo, 1999; Galindo et al., 1999).

23

CÁLCULO RELACIONAL DIFUSO

- Definimos un **Cálculo de Dominios**, porque es más fácil de usar:
 - Es más explícito (al usar variables de dominio).
 - Tiende a estar más cerca del Lenguaje Natural.
- **Formato Expresiones:** $\{x_1, \dots, x_n \mid y(x_1, \dots, x_n)\}$
 - A la izquierda se ponen las **variables (atributos)** que queremos que aparezcan en el resultado.
 - A la derecha aparecerá una **Fórmula Bien Formada (FBF o WFF, Well Formed Formula)**, con esas variables como únicas variables *libres*.
 - Expresa una **condición** que debe satisfacerse por todas las tuplas del resultado.
 - Un predicado sólo puede ser **Verdadero o Falso** (está basado en el Cálculo de Predicados de primer orden).
 - El cálculo relacional difuso no altera esa condición, ya que al final una tupla sólo puede estar en el resultado (aparecer) o no estar (no aparecer).
 - Sin embargo, a las tuplas que aparezcan en el resultado se les podrá dar un grado de pertenencia que indique en qué medida pertenecen a ese resultado.

24

CÁLCULO RELACIONAL DIFUSO

- Las FBF se forman a partir de los llamados **Átomos Difusos**:
 - Formados con un **Grado de Cumplimiento** (Δ) y un **Umbral** (g), con el siguiente formato: $\Delta \supset g$
 - Puede utilizarse también el comparador ">" u otro distinto. En este último caso cambiaría totalmente el significado de la consulta.
 - El **umbral** establece el mínimo valor de verdad de Δ para que sea **Verdad**: Por defecto se supone >0 (con grado no nulo).
 - Grado de cumplimiento** Δ puede ser de dos tipos:
 - Pertenencia**: $\Delta = R(x_1, \dots, x_n) = \max_{r=1, \dots, m} \{ \min_{c=1, \dots, n} \{ \Theta^=(A_{rc}, x_c) \} \}$
 - R es una Relación Difusa de grado n y cardinalidad m , y los A_{rc} son los valores de los atributos de la fila (tupla) r y la columna (atributo) c .
 - La fórmula mide el **grado con el que la tupla** $\{x_1, x_2, \dots, x_n\}$ **pertenece a la relación** R . Dicho de otra forma: indica el grado con el que dicha tupla se parece a una tupla de R , que sea la tupla más parecida a ella.
 - $\Theta^=$ es un comparador difuso de igualdad (por ejemplo la medida de Posibilidad).
 - Comparación**: $\Delta = \Theta^q(x, y)$
 - Θ^q es un comparador difuso inducido por el comparador θ ($>, \geq, < \dots$).
 - x e y son variables de dominio o constantes (crisp o difusas).
 - Estas comparaciones difusas expresan el grado de cumplimiento con el que x e y se relacionan entre sí con el comparador Θ^q (mayor difuso...). 25

CÁLCULO RELACIONAL DIFUSO

- Formación de las FBF (WFF)**: Se construyen de manera muy parecida a las del cálculo relacional clásico:
 - Las variables (de dominio, en este caso), en una **FBF**, pueden ser **LIBRES** (*free*) o **LIGADAS** (*bound*).
 - La mayor novedad es la posibilidad de incluir **Cuantificadores Difusos** (Q): "la mayoría", "pocos"...
 - Formación de las FBF**:
 - 1. Un **átomo** es una **FBF**. Sus variables son libres.
 - 2. Si F es una **FBF**, también lo es **NOT(F)**. Sus variables son libres o ligadas según lo sean en F .
 - 3. Si $F1$ y $F2$ son **FBF**, también lo son **(F1 AND F2)** y **(F1 OR F2)**. Sus variables son libres o ligadas según lo sean en $F1$ y $F2$.
 - 4. Si $F(x)$ es una **FBF** y x una variable libre de $F(x)$, entonces, también son **FBF** $\$x(F(x))$, $" x(F(x))$ y $Qx(F(x))$ (donde Q es un cuantificador difuso). La variable x , estará ligada en las nuevas **FBF**.
 - 5. Nada más es una **FBF**. Pueden incorporarse paréntesis y otros operadores en función de los anteriores: La **implicación** \supset , **XOR** (*eXclusive OR*), **NAND**, **NOR**...

CÁLCULO RELACIONAL DIFUSO

- **Transformaciones en las FBF:** SÓLO son NECESARIOS los operadores: NOT, OR (o AND), \forall (o \exists) y los cuantificadores difusos.
 - **Leyes de DeMorgan:**
 - $F1 \text{ AND } F2 \circ \text{NOT}(\text{NOT}(F1) \text{ OR } \text{NOT}(F2))$
 - $F1 \text{ OR } F2 \circ \text{NOT}(\text{NOT}(F1) \text{ AND } \text{NOT}(F2))$
 - **Operación XOR, (“O Exclusivamente”, eXclusive OR):**
 - $F1 \text{ XOR } F2 \circ (F1 \text{ AND } \text{NOT}(F2)) \text{ OR } (\text{NOT}(F1) \text{ AND } F2)$
 - **Implicación:**
 - $F1 \text{ ® } F2 \circ \text{NOT}(F1) \text{ OR } F2$
 - **Cuantificadores:**
 - $\forall x(F(x)) \circ \text{NOT } \exists x(\text{NOT } F(x))$
 - $\exists x(F(x)) \circ \text{NOT } \forall x(\text{NOT } F(x))$
 - $\forall x(\text{NOT}(F(x))) \not\equiv \text{NOT } \forall x(F(x))$
 - $\exists x(\text{NOT}(F(x))) \not\equiv \text{NOT } \exists x(F(x))$
- **Expresiones seguras (safe):**
 - Son aquellas que producen relaciones finitas.
 - En cálculo relacional clásico las expresiones no seguras se descartan pues carecen de sentido. Por ejemplo: $\{x_1, \dots, x_n \mid \neg R(x_1, \dots, x_n)\}$
 - En cálculo relacional difuso las expresiones no seguras desde un punto de vista estricto pueden ser útiles, por lo que se aplica una “**evaluación limitada**”, restringiendo la evaluación de la consulta a valores coherentes (contenidos y relacionados en alguna relación de la FBF).

27

CÁLCULO DIFUSO: Operadores Algebraicos

- **Capacidad Expresiva:** Está demostrado que cualquier expresión del Álgebra Relacional Difuso se puede expresar en este Cálculo Rel. Difuso.
- **Operadores PRIMITIVOS del Álgebra:** (NOT = \neg , AND = $\hat{\cup}$, OR = $\hat{\cup}$)
 - **Unión:** $R \hat{\cup} S = \{x_1, \dots, x_n \mid R(x_1, \dots, x_n) \vee S(x_1, \dots, x_n)\}$
 - No tiene sentido establecer umbrales de cumplimiento en ninguna de las dos relaciones difusas, ya que las tuplas del resultado deben estar en alguna de las dos relaciones (o en ambas).
 - Esa expresión no es segura desde un punto de vista estricto.
 - Podemos “inventarnos” una tupla que no exista ni en R ni en S y que exista una tupla muy parecida a ella en R o en S (con grado 1).
 - De ahí la necesidad de aplicar la “**evaluación limitada**”.
 - **Diferencia:** $R - S = \{x_1, \dots, x_n \mid R(x_1, \dots, x_n) \wedge \neg S(x_1, \dots, x_n) \geq g\}$
 - Tuplas de R que no están en S con un grado mayor o igual a g .
 - **Producto Cartesiano:**

$$R \times S = \{x_1, \dots, x_n, v_1, \dots, v_m \mid R(x_1, \dots, x_n) \wedge S(v_1, \dots, v_m)\}$$
 - **Proyección:**

$$\Pi_{A_1, \dots, A_k}(R) = \{x_1, \dots, x_k \mid \exists x_{k+1}, \dots, x_n R(x_1, \dots, x_n)\}$$

28

CÁLCULO DI FUSO: Operadores Álgebraicos

- **Selección:** $S_C(R) = \{x_1, \dots, x_n \mid R(x_1, \dots, x_n) \wedge C'\}$
 - C' es la condición C cambiando los atributos de R por sus variables de dominio correspondientes (los x_i).
- **Operadores NO PRIMITIVOS del Álgebra:**
 - **Intersección:** Tuplas que pertenecen a R con grado mayor o igual que γ y que pertenecen a S con grado mayor o igual que γ' :

$$R \cap S = \{x_1, \dots, x_n \mid R(x_1, \dots, x_n) \geq g \wedge S(x_1, \dots, x_n) \geq g'\}$$
 - Esta expresión no sería segura desde un punto de vista estricto.
 - **Reunión:** La condición C debe relacionar atributos de R con los de S :

$$R \bowtie_C S = \{x_1, \dots, x_n, v_1, \dots, v_m \mid R(x_1, \dots, x_n) \wedge S(v_1, \dots, v_m) \wedge C'\}$$
 - **División:** Obtiene las tuplas $A=\{a_1, \dots, a_n\}$ tal que si $B=\{b_1, \dots, b_m\}$ pertenece a S , entonces (A,B) pertenece a R en grado mínimo γ .

$$R \div S = \{a_1, \dots, a_n \mid S(b_1, \dots, b_m) \rightarrow R(a_1, \dots, a_n, b_1, \dots, b_m) \geq g\}$$
- Con las ecuaciones anteriores es fácil **traducir** una expresión en **Álgebra Relacional Difusa** a una expresión en **Cálculo Relacional Difuso**.

29

CÁLCULO DI FUSO: Ejemplos

- **Ejemplos:**
 - “Mostrar los datos de los jugadores Buenos (en grado mínimo 0.5)”:

$$\{j, e, a \mid \exists c (R(j, e, a, c) \wedge \Theta^-(c, \mathbf{Bueno}) \geq 0.5)\}$$
 - “Mostrar los jugadores con sus equipos y alturas de los que sean de Jaén o Málaga con altura Normal (en grado mínimo 0.5) o Bajos (en grado mínimo 0.7)”:

$$\{j, e, a \mid \exists c (R(j, e, a, c) \wedge (e = \mathbf{Jaén} \vee e = \mathbf{Málaga}) \wedge (\Theta^-(a, \mathbf{Normal}) \geq 0.5 \vee \Theta^-(a, \mathbf{Bajo}) \geq 0.7))\}$$
 - “Mostrar los equipos en los que TODOS sus jugadores sean Bajos (en grado mínimo 0.5) o Buenos (en grado mínimo 0.75)”:

$$\{e \mid \forall j, a, c (R(j, e, a, c) \rightarrow (\Theta^-(a, \mathbf{Bajo}) \geq 0.5 \vee \Theta^-(c, \mathbf{Bueno}) \geq 0.75))\}$$

30

CÁLCULO DIFUSO: Función D

- **Relación Resultante de una expresión en Cálculo:**
 - 1. **Calcular la componente de valor:** Las tuplas que cumplen el predicado propuesto.
 - 2. **Calcular la componente de compatibilidad (atributos o grados de compatibilidad de la Relación Difusa Resultante):** Calcular el **grado** con el que cada valor de cada tupla del resultado satisface la condición impuesta en la expresión del Cálculo Difuso. Para esto se usa la función $D_x^S(y)$
- **Función $D_x^S(y)$, o “Grado de una variable x en una FBF y con una sustitución S ”:**
 - Para cada tupla S del resultado y cada variable x de la expresión en Cálculo Difuso, se **calcula el grado de cumplimiento de la FBF y** por parte de esa variable x : $D_x^S(y) \in [0,1]$ donde 1 es una constante que indica que ese grado no es aplicable (por ejemplo si la variable no aparece en y): Suponemos $1 < 0$.
 - Utilizando sólo los operadores básicos se pueden dar **cuatro casos**: Sin operadores (y es un átomo de Pertenencia o de Comparación), y que el operador principal de y sea la negación, la disyunción o el cuantificador existencial (\exists): La definición es recursiva con base en el caso 1.

31

CÁLCULO DIFUSO: Función D

- 1. **Sin operadores** (y es un átomo):
 - **Átomo de Pertenencia R :** Si K es una lista con las constantes que aparecen en R :

$$\Delta_x^S(R(x_1, \dots, x_n, K) \geq g) = \begin{cases} R(K) & \text{si no hay variables en } R \\ R(S, K) & \text{si } x = A_i \text{ y } \nexists C_i \\ \min\{c_{ri}, R(S, K)\} & \text{si } x = A_i \text{ y } \exists C_i \\ 1 & \text{en otro caso} \end{cases}$$
 donde c_{ri} es el valor del atributo de compatibilidad de la tupla más parecida a (S, K) en R (si existen varios se tomará el mayor).
 - **Átomo de Comparación $Q^g(x_i, y) \geq g$:**

$$\Delta_x^S(Q^g(x_i, y) \geq g) = \begin{cases} \Theta^g(s_i, y) & \text{si } x_i \text{ es una variable y } x = x_i \\ \Theta^g(x_i, y) & \text{si } x_i \text{ es una constante de la variable } x \\ 1 & \text{en otro caso} \end{cases}$$
 - Observe que no se usa el umbral g de cada átomo: Ese umbral sirve para discernir si una tupla pertenece o no al resultado (cumple o no cumple la condición), pero una vez que la tupla pertenece al resultado no se necesita para ver en qué medida ese valor de esa tupla pertenece al resultado.

32

CÁLCULO DI FUSO: Función D

- **2. Negación:** $y(x_1, x_2, \dots, x_n) = \neg y_1(x_1, x_2, \dots, x_n)$

$$\Delta_x^S(y(x_1, \dots, x_n)) = \begin{cases} 1 - \Delta_x^S(y_1(x_1, \dots, x_n)) & \text{si } \Delta_x^S(y_1(x_1, \dots, x_n)) \neq 1 \\ 1 & \text{si } \Delta_x^S(y_1(x_1, \dots, x_n)) = 1 \end{cases}$$

- **3. Disyunción:** $y(x_1, x_2, \dots, x_n) = y_1(x_1, x_2, \dots, x_n) \text{ OR } y_2(x_1, x_2, \dots, x_n)$

$$\Delta_x^S(y(x_1, \dots, x_n)) = \max\{\Delta_x^S(y_1(x_1, \dots, x_n)), \Delta_x^S(y_2(x_1, \dots, x_n))\}$$

– Sale simple porque hemos supuesto $1 \leq 0$.

- **4. Cuantificador Existencial:** $y(x_1, x_2, \dots, x_n) = \exists x_{n+1} y_1(x_1, x_2, \dots, x_n)$

$$\Delta_x^S(y(x_1, \dots, x_n)) = \max_{s_{n+1} \in \text{DOM}(y)} \{\Delta_x^S(y_1(x_1, \dots, x_n, s_{n+1}))\}$$

– $\text{DOM}(y)$ restringe el ámbito de actuación de la variable cuantificada.

- **Averiguar el grado con el que toda la tupla, globalmente, satisface la FBF y :**

- Es fácil extender esta definición para el caso en el que no nos refiramos a una variable x concreta
- Hay que modificar sólo el caso base (caso 1).

33

CÁLCULO DI FUSO: Cuantificadores Difusos

- **Formato General** con el Cuantificador Q , con $\rho, \gamma \in [0, 1]$: (Galindo et al., 2000a)

$$Q_g^r x_{n+1} y(x_1, \dots, x_n)$$

- **Ejemplo:** “Equipos en los que casi todos (con grado mínimo 0 y $\rho=0.8$) sus jugadores son Buenos (con grado mínimo 0.75)”:

$$\{e \mid \text{Casi_todos}_{0.8} j, a, c (R(j, e, a, c) \rightarrow \Theta^-(c, \text{Bueno}) \geq 0.75)\}$$

- **Evaluación del Cuantificador Difuso:** Aspectos distintos a considerar:

- **T:** Total de elementos que cumplen la condición de la FBF y sobre la que influye el cuantificador Q .
 - Si Q es relativo se divide por el total de elementos existentes.
- **M:** Suma de los grados de cumplimiento de los elementos que cumplen la condición de la FBF y sobre la que influye Q .
 - Si Q es relativo se divide por el total de elementos existentes.

- El parámetro r (superíndice) mide la importancia de T con respecto a M .
- Finalmente, el grado de cumplimiento de cualquier variable x en la FBF y con la sustitución (tupla) S es:

$$r T + (1-r) M \Rightarrow \Delta_x^S(y) = Q(r T + (1-r) M)$$

34

CÁLCULO DI FUSO: Cuantificadores Difusos

Ejemplo:

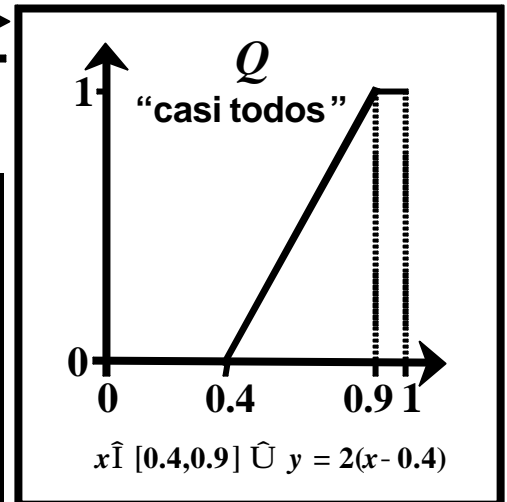
- “Equipos en los que **casi todos** (con grado mínimo $\gamma = 0$ y $r = 0.8$) sus jugadores son Buenos (con grado mínimo 0.75)”:

$$\{e \mid \text{Casi_todos}_{0.8} j, a, c (R(j, e, a, c) \rightarrow \Theta^-(c, \text{Bueno}) \geq 0.75)\}$$

- Supongamos los siguientes valores de T y M :

EQUIPO	T	M	$\rho T + (1-\rho)M$	C_{EQUIPO}
Almería	1/2	0.75/2	0.475	0.15
Cádiz	2/2	1.75/2	0.975	1.00
Córdoba	1/2	0.75/2	0.475	0.15
Granada	0/3	0/3	0.000	0.00
Jaén	3/4	2.25/4	0.713	0.63
Málaga	2/3	1.5/3	0.633	0.47
Sevilla	3/3	3/3	1.000	1.00

Cuantificador Difuso Relativo:



35

Bibliografía

- J. Baldwin, "Knowledge Engineering Using a Fuzzy Relational Inference Language". Proc. IFAC Symp. on Fuzzy Information Knowledge Representation and Decision Analysis, pp. 15-21, 1983.
- D. Barbara, H. Garcia-Molina, D. Porter, "The Management of Probabilistic Data". IEEE Trans. on Knowledge and Data Engineering, 4, pp. 487-502, 1992.
- P. Bosc, M. Galibourg, G. Hamon, "Fuzzy Querying with SQL: Extensions and Implementation Aspects". Fuzzy Sets and Systems, 28, pp. 333-349, 1988.
- P. Bosc, O. Pivert, K. Farquhar, "Integrating Fuzzy Queries into an Existing Database Management System: An Example". International Journal of Intelligent Systems, 9, pp. 475-492, 1994.
- P. Bosc, O. Pivert, "SQLf: A Relational Database Language for Fuzzy Querying". IEEE Transactions on Fuzzy Systems, 3, pp. 1-17, 1995.
- P. Bosc, D. Dubois, O. Pivert, H. Prade, "Flexible queries in relational databases -- The example of the division operator". Theoretical Computer Science 171, pp. 281-302, 1997.
- B.P. Buckles, F.E. Petry, "A Fuzzy Representation of Data for Relational Databases". Fuzzy Sets and Systems, 7, pp. 213-226, 1982.
- B.P. Buckles, F.E. Petry, "Extending the Fuzzy Database with Fuzzy Numbers". Information Sciences 34, pp. 45-55, 1984.
- B.P. Buckles, F.E. Petry, H.S. Sachar, "A Domain Calculus for Fuzzy Relational Databases". Fuzzy Sets and Systems, 29, pp. 327-340, 1989.
- R. Cavallo, M. Pittarelli, "The Theory of Probabilistic Databases". Proc. 3th International Conference on Very Large Databases, pp. 102--110, 1987.

36

Bibliografía

- E.F. Codd, "A Relational Model of Data for Large Shared Data Banks". Communications of the ACM, 13(6), pp. 377-387, June 1970.
- E.F. Codd, "Relational Completeness of Data Base Sublanguages". Data Base Systems. Courant Computer Science Symposia Series, Vol. 6, pp. 65-98, Englewood Cliffs, Nueva Jersey. Prentice Hall, NY 1972.
- E.F. Codd, "Extending the Database Relational Model to Capture More Meaning". ACM Trans. on Database Systems, 4, pp. 262-296, 1979.
- E.F. Codd, "Missing Information (Applicable and Inapplicable) in Relational Databases". ACM SIGMOD Record, Vol. 15(4), 1986.
- E.F. Codd, "The Twelve Rules for Relational DBMS". San José, The Relational Institute, Technical Report EFC-6, 1986.
- E.F. Codd, "More Commentary on Missing Information in Relational Databases". ACM SIGMOD Record, Vol. 16(1), 1987.
- E.F. Codd, "The Relational Model for Database Management, Versión 2". Reading Mass.: Addison-Wesley, 1990.
- C.J. Date, "Null Values in Database Management". In "Relational Database: Selected Writings", Ed. C.J. Date. Reading Mass.: Addison-Wesley, 1986.
- R. Elmasri, S.B. Navathe, "Fundamentals of Database Systems", Third Edit. Addison-Wesley, 2000.
- N. Fuhr, "A Probabilistic Framework for Vague Queries and Imprecise Information in Databases". Proc. 6th International Conf. on Very Large Data Bases, pp. 77--85, 1990.

37

Bibliografía

- J. Galindo, J.M. Medina, O. Pons, J.C. Cubero, "A Server for Fuzzy SQL Queries". In "Flexible Query Answering Systems", eds. T. Andreassen, H. Christiansen and H.L. Larsen, Lecture Notes in Artificial Intelligence (LNAI) 1495, pp. 164-174. Ed. Springer, 1998a.
- J. Galindo, J.M. Medina, A. Vila, O. Pons, "Fuzzy Comparators for Flexible Queries to Databases". Iberoamerican Conference on Artificial Intelligence, IBERAMIA'98, Lisbon (Portugal), pp. 29-41, October 1998b.
- J. Galindo, "Tratamiento de la Imprecisión en Bases de Datos Relacionales: Extensión del Modelo y Adaptación de los SGBD Actuales". Ph. Doctoral Thesis, University of Granada (Spain), March 1999 (www.lcc.uma.es).
- J. Galindo, J.M. Medina, M.C. Aranda Garrido, "Querying Fuzzy Relational Databases Through Fuzzy Domain Calculus". International Journal of Intelligent Systems, Vol. 14(4), pp. 375-411, 1999.
- J. Galindo, J.M. Medina, J.C. Cubero, M.T. García, "Fuzzy Quantifiers in Fuzzy Domain Calculus". 8th Int. Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU'2000, pp. 1697-1704. Madrid (Spain), Julio 2000a (www.lcc.uma.es).
- J. Galindo, J.M. Medina, J.M. Rodríguez, "Comparadores para Bases de Datos Difusas: Definiciones, Clases y Relaciones". X Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF'2000), pp. 187-192. Sevilla (Spain), September 2000b (www.lcc.uma.es).
- J. Grant, "Incomplete Information in a Relational Database". Fundamenta Informaticae, 3, pp. 363-378, 1980.
- J.M. Medina, O. Pons, M.A. Vila, "GEFRED. A Generalized Model of Fuzzy Relational Databases". Information Sciences, 76(1-2), 87-109, 1994.
- J.M. Medina, "Bases de Datos Relacionales Difusas. Modelo Teórico y Aspectos de su Implementación". Ph. Doctoral Thesis, University of Granada (Spain), 1994 (www.decsai.ugr.es).

38

Bibliografía

- N. Mouaddib , ‘Fuzzy Identification in Fuzzy Databases: The Nuanced Relational Division’. International Journal of Intelligent Systems, 9, pp. 461-473, 1994.
- F.E. Petry, “Fuzzy Databases: Principles and Applications” (with chapter contribution by Patrick Bosc). Int. Series in Intelligent Technologies. Ed. Zimmermann. Kluwer Academic Publishers, 1996.
- H. Prade, C. Testemale, “Generalizing Database Relational Algebra for the Treatment of Incomplete/Uncertain Information and Vague Queries”. Informat. Sciences 34, pp. 115-143, 1984.
- H. Prade, C. Testemale, “Fuzzy Relational Databases: Representational issues and Reduction Using Similarity Measures”. J.Am. Soc. Information Sciences 38(2), pp. 118-126, 1987.
- S. Sheno, A. Melton, “An Extended Version of the Fuzzy Relational Database Model”. Information Sciences, 51, pp. 35-52, 1990.
- A. Shoshani, H. Wong, “Statistical and Scientific Database Issues”. IEEE Trans. Software Engineering, 11, pp. 235-256, 1985.
- V. Tahani, “A Conceptual Framework for Fuzzy Query Processing--A Step toward Very Intelligent Database Systems”. Information Processing and Management, 13, pp. 289-303, 1977.
- M. Umamo, “Freedom-O: A Fuzzy Database System”. In “Fuzzy Information and Decision Processes”. Eds. M. Gupta , E. Sanchez North-Holland, Amsterdam, Pub. Comp., pp. 339-347, 1982.
- M. Umamo, S. Fukami, “Fuzzy Relational Algebra for Possibility-Distribution-Fuzzy-Relational Model of Fuzzy Data”. Journal of Intelligent Information Systems, 3, pp. 7-28, 1994.
- M. Wong, K. Leung, “A Fuzzy Database-Query Language”. Inform. Systems 15, pp. 583-590, 1990.
- M. Zemankova, A. Kandel, “Fuzzy Relational Databases - A Key to Expert Systems”. Köln, Germany, TÜV Rheinland, 1984.
- M. Zemankova, A. Kandel, “Implementing Imprecision in Information Systems”. Information Sciences, 37, pp. 107-141, 1985.